

Glioma Segmentation and Classification System Based on Proposed Texture Features Extraction Method and Hybrid Ensemble Learning



Kirti Raj Bhatele^{1*}, Sarita Singh Bhadauria²

¹ CSE & IT, Rajiv Gandhi Technical University, Bhopal 462033, India

² Madhav Institute of Science & Technology, Gwalior 474005, India

Corresponding Author Email: kirtirajbhatele8@gmail.com

<https://doi.org/10.18280/ts.370611>

ABSTRACT

Received: 3 March 2020

Accepted: 23 November 2020

Keywords:

Thresholding, High Grade Glioma (HGG), Low Grade Glioma (LGG), DWT (Discrete wavelet transform), LBP (Local Binary pattern), GLRLM (Gray level run length Matrix) Enhanced wavelet binary pattern run length matrix method (EWBPRL), XGBoost with Random forest (XGBRF)

This paper presents an efficient and accurate automated system based on the hybrid XGBoost with Random forest (XGBRF) ensemble model in order to classify the Glioma (type of mostly diagnosed brain tumor) into low grade and high grade Glioma. In this approach initially global thresholding is employed on various MRI sequence and their fusion combinations in order to perform the accurate segmentation. Then uses a proposed Enhanced wavelet binary pattern run length matrix method (EWBPRL) for textural features extraction from the region of interest or segmented Glioma tumor region. This proposed feature extraction method is based on the Discrete wavelet transform (DWT), Local Binary pattern (LBP) and Gray level run length Matrix (GLRLM) methods to extract texture features from the segmented region. Some morphological features are also computed from the segmented region along the textural features. Finally both these extracted features are employed in order to train a hybrid XGBoost with Random forest ensemble model for the first time. The proposed automated system apart from accurately detecting and segmenting the tumors region from the fused MRI sequences, also tends to determine the grading of Glioma in terms of severity. The proposed system is evaluated on the large size balance local dataset and as well as on the popular global datasets like BRATS 2013 and BRATS 2015. This approach offers an encouraging accuracy of 99.25% on the local dataset with the fusion of T1C+T2+Flair MRI sequence as compare to 96.75% accuracy, which is achieved utilizing the fusion of T1+T1C+T2+Flair MRI sequence.

1. INTRODUCTION

As per the World health organization reports and other health related surveys are showing that brain tumors patients are increasing on a global level. Especially some of the reasons behind this increasing number of patients suffering from the brain tumor might be the radiations from devices like mobile etc., unhealthy eating's, over-thinking, stressed life, and many more factors. As per the report around seventy eight thousand cases of patients suffering from brain tumors to be diagnosed at the end of the year 2016 is published by Brain Tumor Association of USA [1]. In India every year around 40,000 to 50,000 death cases are reported because of brain tumors and out of those around 20% are of children as per the report published by Brain Tumor Foundation of India. In fact it is to be believed that around ninety percent of these cases can be treated and cured only if the tumors are diagnosed at the early stage and given proper treatment [2].

A brain tumor is a lump in the brain which is a result of abnormal growth of brain tissues in brain. The so called intrinsic characteristics of brain tissues are responsible distinguishing brain tumors from each other [3]. When it comes to the diagnosis of brain tumors, it's the MRI scans modalities namely T1, T2, Flair, T1 Contrast etc. which are mostly used by the radiologists and researchers for the detection and classification of brain tumors. Primary and secondary are the two main broad categories of brain tumors

[3]. When tumors originate and evolved in brain tissues itself is known as primary brain tumor. The secondary brain tumors are those, which originates and discovered in other parts of the body and then extend towards brain cells. Glioma tumor is the mostly diagnosed primary tumor and as it mostly originates on the glial cells of brain so it is known as the Glioma. The word Glioma encompasses many different tumor types, and Gliomas come in different "grades." Gliomas are often categorized as "low-grade", meaning that the tumor cells look as if they are dividing more slowly under the microscope, or "high-grade," meaning that the cells look more aggressive under the microscope. Low-grade tumors can cause problems even when they are not malignant (cancerous) by pressing on the normal structures in the brain and causing symptoms. Gliomas develop from the abnormal growth of glial cells. A glial cell is a type of cell in the brain that surrounds the neurons, which are the cells that conduct electrical impulses to transmit information. Glial cells support and insulate the neurons. For example, an astrocyte is a special form of glial cell. Astrocytomas are a type of Glioma that forms because of abnormal growth of astrocytes [4].

There is already a lot of research carried out in the field of effective and accurate brain tumor segmentation using conventional image processing approaches but very little efforts have been made in the direction of the further classification of brain tumor [5]. Apart from this, the datasets used for validation in most of the approaches are of small size

and unbalance hence giving insufficient results. All these factors simply restrict the use of such system developed by the researchers to be used in real life by the radiologist. Considering all these facts, the proposed system is focused on the development of an efficient and robust approach employed for the segmentation and classification of Glioma into low and high grade types. This approach is tested and evaluated on a balance local dataset as well as on the global datasets like BRATS (Brain tumor Segmentation challenge) 2013 and 2015. This paper consists of six sections. The first and second section deals with the introduction and literature review of the existing approaches in this field. Then third section simply illustrates the proposed work with brief description. Result and evaluation is presented in the fourth section followed by a discussion in the fifth section. Lastly the conclusion is given in the sixth section.

2. LITERATURE REVIEW

Since various semi-automatic and automatic approaches have been already proposed for the segmentation of brain tumor from the MRI sequence. Generally there are four MRI imaging sequence used for the diagnosis of brain tumor namely T1-weighted (T1w), T2-weighted (T2w), T1-postcontrast (T1C) and Fluid Attenuated Inversion Recovery (Flair). In most of the segmentation approaches, T2 MRI modality is used. Although due to the complex structure of human brain and anatomy, all the four MRI modalities are used by the radiologist to diagnose the tumor and then classify the type of brain tumor. The initial segmentation approaches are majorly based on general methods like Thresholding, Region Growing, Fuzzy based, Brain symmetry based etc. In 2000, an approach proposed by Boudraa et al. [6] for the detection of anomaly (tumor) from the MRI sequence in which local contrast enhancement as preprocessing then followed by the fuzzy-c-means clustering is employed for accurate segmentation. This approach achieved accuracy of 88% on a dataset of size of ten patients. Where as in 2001, Van Leemput et al. [7] proposed a stochastic tissue intensity model based on contextual information using Markov Random field to segment the brain tumor from the MRI sequences. In 2004, a tumor detection method using histograms was proposed by Wang et al. [8]. In 2007, Ray et al. [9] proposed a technique for computing bounding boxes around brain abnormality in standard MR images based on symmetry. The technique uses a scoring function that provides a measure of the similarity or difference between two regions in terms of the Bhattacharya coefficient computed on those region intensity histograms. In 2008, Khayati et al. [10] come up with a model based on Markov Random Field and adaptive filters in order to segment lesion from the Flair MRI sequence. In 2009, Wang and Fei [11] devise an approach by incorporating modified FCM classifier and multi scale diffusion filtering to segment the brain tumour accurately.

In 2010, Yamamota et al. [12] proposed an approach for the lesion detection composed of a level set method, rule-based method and a support vector machine. This method was applied 49 slices selected from 6 studies of three MS cases including 168 MS lesions. The resulting sensitivity for detection of lesions was 81.5%. In 2013, Sharma and Mukherjee [13] proposed an approach which uses FCM for tumor segmentation and contour method to obtain the boundary of the brain tumor precisely. Then Adaptive network

fuzzy inference system and genetic algorithm are employed for the selection of textural features. In 2014 El-Dashan et al. [14, 15] come up with a hybrid approach in which first the features are extracted by discrete wavelet transform (DWT) and then PCA is employed. The k-nearest neighbor and feed-forward back-propagation artificial neural network classifiers are used for the classification and hence achieving accuracy of 98% and 97% respectively. In 2013, Arakeri and Reddy [16] come up with a computer aided detection system (CAD) to classify tumour as malignant and benign. In this system, wavelet transform and Advance FCM is used to segment the anomaly region and then textural and shape features are extracted from the segmented region. These features are used to train the ensemble classifier and achieved an accuracy of 99%. In 2015, Gupta and Khanna [17] also proposed a CAD system to detect lesion based on Block wise segmentation which resulted in four blocks. Texture and shape based features are extracted and calculated from these four blocks. This CAD system achieved an accuracy of 97.93% on a dataset of size of 650 MRIs. Another method in the same year was proposed by Subashini et al. [18]. This method used FCM and extracted features are shape and textural based. These features are used to train a Naive Bayes classifier and claimed accuracy of 91%. In 2016, Vishuvarthanan et al. [19] developed an automatic clustering method for effective segmentation of brain tumour. This method consisted of hybrid self-organizing map method and Fuzzy k means. The sensitivity of this method was 97.37%. In 2017, Zhan et al. come up with an automated approach for the grading of Glioma with the help of fusion of intensity, volume and texture features along with KNN classifier to deliver accuracy of 87.59 on the BRATS 2015 dataset [20]. In 2018, Gupta et al. [21] proposed a clinical decision support system in order to assist the radiologist. This system utilizes the fusion of MRI pulse sequences and then applies adaptive thresholding for effective segmentation. The features are extracted by using LBP and Run length of centralized pattern. Since it is used for the classification of Glioma i.e. Multiclass identification so Naive Bayes classifier is used and achieved an accuracy of 97.83%. In 2019, Gupta et al. [22] developed an approach in which features from the segmented region is extracted with the help of the Fusion of LBP, GLCM and GLRL methods and then these features are used for the classification of Glioma with the help of ensemble learning.

From the literature review, it can be concluded that the accuracy of all these approaches ranges from 70% to 98%. Most of the existing approaches have used small size and unbalanced dataset, which is not good in terms of getting reliable results and real time application. Although there is still scope for improvement in the areas related to effective and accurate segmentation of brain tumor from the MRI sequences. Then followed by the new and more effective features extraction methods. Finally experimentation related to the selection of correct machine learning and ensemble learning models for delivering more accurate results. So in order to propose an effective and accurate system for the classification of Glioma types in real time. This work suggests a MRI sequence fusion based segmentation method and then from the segmented region textural features are extracted with the help of our proposed Enhanced wavelet binary pattern run length matrix method along with the morphological features. Then a hybrid ensemble model i.e. XGBoost with random forest (XGBRF) is employed first time for the classification of brain tumors or any cancer type.

3. PROPOSED APPROACH

The proposed approach or system consist of four major stages. The first stage is all about doing the preprocessing and resizing of the MRI sequence images. Then fusion and segmentation of Glioma brain tumor from the fused MRI sequences like T1C+T2+Flair MRI and T1+T1C+T2+Flair are done utilizing the global thresholding segmentation method. The third stage is all about the textural features and morphological or shape based features calculation from the segmented Glioma region. The textural features are extracted employing the proposed EWBPRL method. This method is based on discrete wavelet transform (DWT), Local binary pattern (LBP) along with Grey level run length matrix (GLRL). The last stage is all about training the hybrid ensemble classifier i.e. XGBRF classifier with the help of the extracted features from the earlier stage. Finally the proposed approach is tested and evaluated on local dataset as well as on the two global datasets i.e. BRATS 2013 and 2015 in order to obtain the results. The overall proposed automated system for the classification of Glioma brain tumour is illustrated with the help of Figure 1 below: (a) Glioma MRI sequences preprocessing and resizing stage, (b) MRI sequences fusion and segmentation stage resulted in segmented Glioma region or region of interest (ROI), (c) Feature extraction stage in which morphological features are extracted directly from the ROI and textural features are extracted using proposed

Enhanced wavelet binary pattern run length matrix (EWBPRL) method, (d) Training and testing stage of proposed automated Glioma classification system using XGBoost with random forest ensemble classifier model.

3.1 Pre-processing and Glioma segmentation

As the Magnetic resonance imaging (MRI) technique is one of the most popular medical imaging techniques employed for the accurate detection and classification of Glioma type by the radiologist. The MRI medical imaging mainly consist of four pulse imaging sequence like T1 weighted (T1), T2 weighted (T2), T1 post contrast weighted (T1C) and Fluid attenuated inversion recovery (Flair). These different MRI pulses imaging sequence renders different image intensity information about the brain region and brain tumour. When these different MRI pulses sequence are fused together then the resulted fused MRI sequence provides more morphological and texture information about the Glioma as compare to the single MRI sequence. The fused MRI sequence also provides clear and precise tumour region which can be easily segregated with the help of the thresholding based segmentation method. The MRI fusions combinations of T2+T1C+Flair and T2+T1+T1C+Flair MRI sequences provides better segmentation results as compare to solo MRI sequence and other fusion combination sequence as it is proved in the result section.

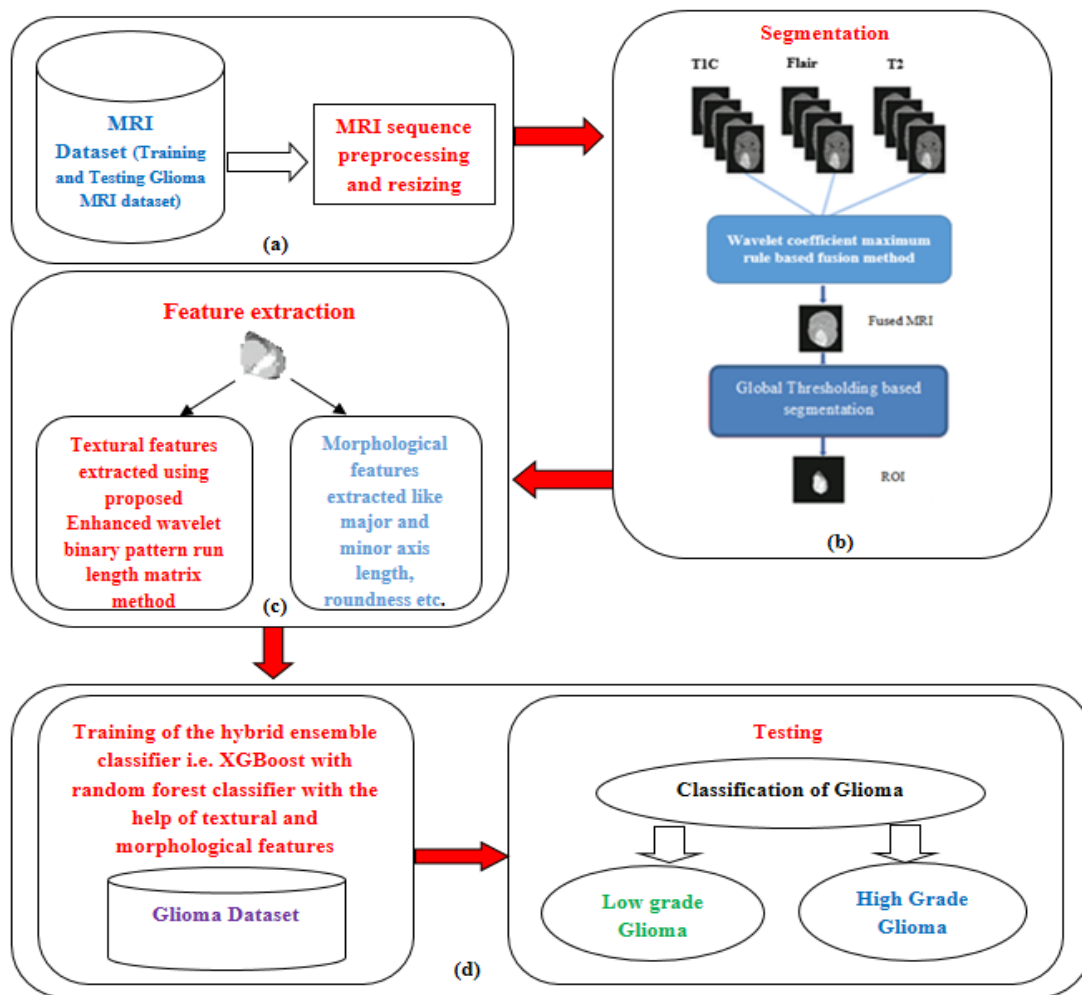


Figure 1. The overall proposed automated system for the classification of Glioma Brain Tumour

In this stage, all the MRI sequences initially pre-processed with the help of antistrophic diffusion filter technique [23] in order to perform enhancement and removal of noise. Then skull stripping is done in order to remove the grey matter from the skull for the sake of effective segmentation of Glioma tumour region. Then fused MRI sequences i.e. (T1C, T1, T2 and Flair), (T1C, T2 and Flair) etc. are obtained employing wavelet coefficient maximum rule based fusion method. Then global thresholding method is employed on the fused MRI sequence to obtain the segmented Glioma brain tumour region [24]. Thresholding methods always delivers good results for brain tumour segmentation. The threshold value used to segregate the Glioma tumour from the fused MRI sequence is simply defined in Eq. (1) below.

$$255 > \text{segmentation threshold} > (\sigma + \mu) \quad (1)$$

where, σ and μ are the standard deviation and mean of fused MRI sequences. The contrast and brightness of the fused MRI sequence is represented by both these factors, so their adequate value will eventually help in getting the accurate segmented region. The threshold value of 45 is delivering the best results in terms of accurate segmentation. The algorithm of this stage is as follows:

Input: Four or three MRI sequence i.e. T1, T1C, T2 and Flair

Output: Segmented Glioma tumour region

Steps:

1. Antistrophic diffusion filter technique is employed on the four T1, T1C, T2 and Flair MRI sequences for image enhancement and noise removal.
2. Resize the MRI sequence as in size of 512*512.
3. Skull striping is carried out in order to segregate the skull from the brain grey matter.
4. MRI sequence combinations is obtained by employing wavelet coefficient maximum rule based fusion method.
5. Global thresholding is applied in order to get the Glioma tumour region from the fused MRI sequences.
6. Finally for precise tumour isolation, applied erosion and dilation.

3.2 Proposed EWPRL method for texture feature extraction

In this stage, two types of features are extracted from the segmented Glioma tumour region which are morphological and textural features. As morphology plays a very important role especially when it comes to the classification of Glioma into low and high grade. The shape of the low grade Glioma is more regular as compare to the shape of the high grade Glioma which is more irregular due to its aggressive spreading [25]. Major and Minor axis length, Roundness, Eccentricity and Equiv Diameter are the five morphological features which are extracted from the segmented Glioma tumour region [26].

Apart from the morphological features, textural features are also very important when it comes to the classification of Glioma [18-22]. As in MRI sequences, cells of brain offer a very powerful textural property. This fact simply allows texture based feature extraction methods play a very significant role in the classification of Glioma [27]. Some of the popular texture feature extraction methods are the Gabor

transform based, discrete wavelet transform (DWT) based, Local binary pattern (LBP), Grey level run length matrix (GLRL) etc. Gabor and Discrete wavelet transform based methods are popular as well as powerful methods. Similarly Local binary pattern is also very popular as it performs the correct characterization of spatial structure of MRI texture based on the extracted binary pattern of neighbourhood pixels [28]. GLRL is also very powerful texture feature extraction method as it is very sensitive to intensity patterns present in the MRI sequence [29]. In order to harness the advantages of all these three popular texture feature extraction methods, a new hybrid feature extraction method is proposed known as the Enhanced wavelet binary pattern run length matrix method (EWPRL) as it is based on the DWT, LBP and GLRL methods. The fusion of DWT along with LBP and GLRL offers feature extraction at multi-resolution i.e. features that might not extracted at one particular resolution now will get extracted at other resolution. This makes the proposed method more robust and efficient. The proposed hybrid feature extraction method is disused below as:

Discrete wavelet transform: The quotient of multi-resolution representation of fused MRI sequence is provided by the DWT. This method helps in analysing fused MRI sequence variations at different resolutions in both time and frequency domain [30]. DWT simply process the segmented Glioma region into the detailed coefficients and approximate coefficients. First level DWT is performed on the segmented Glioma region, which results into four sub bands namely LL (low-low), HH (high-high), HL (high-low), and LH (low-high). The DWT of segmented Glioma region as $f(a,b)$ of size $R*S$ is defined by Eq. (2) and (3) below.

$$W_{\varphi}(j_0, r, s) = \frac{1}{\sqrt{RS}} \sum_{a=0}^{R-1} \sum_{b=0}^{S-1} f(a, b) \varphi_{j_0, r, s}(a, b) \quad (2)$$

$$W_{\psi}^i(j, r, s) = \frac{1}{\sqrt{RS}} \sum_{a=0}^{R-1} \sum_{b=0}^{S-1} f(a, b) \psi_{j, r, s}^i(a, b) \quad (3)$$

where, φ and ψ represents the wavelets and scaling function, $W_{\psi}^i(j, r, s)$ are the approximation coefficients off (x, y) at scale j_0 and j_0 is any arbitrary starting scale. The $W_{\varphi}(j_0, r, s)$ coefficients add horizontal, vertical, and diagonal details for scales $j \geq j_0$. Out of these resultant four sub-band images of the segmented Glioma region as $f(a,b)$, three sub-band images LH, HL, HH are the detail images along horizontal, vertical and diagonal directions respectively. The LL sub-band is the approximation image. The HH sub-band is discarded as it mostly contains noise [31]. The mathematical equations of approximation (LL), horizontal (HL), Vertical (LH) and diagonal (HH) coefficients or sub bands of the segmented Glioma region $f(a,b)$ are defined by Eq. (4), Eq. (5), Eq. (6), Eq. (7) below:

$$\varphi(a, b) = \varphi(a) \varphi(b) \quad (4)$$

$$\psi^H(a, b) = \psi(a) \varphi(b) \quad (5)$$

$$\psi^V(a, b) = \varphi(a) \psi(b) \quad (6)$$

$$\psi^D(a, b) = \psi(a) \psi(b) \quad (7)$$

Table 1. Description of texture based features extracted using proposed EWBPRM method and morphological features

S.No	Texture features calculated using LBP+GLRLM	Morphological features
1	Short Run Emphasis (SRE) measures the distribution of short runs: $SRE = \frac{1}{n_r n_c} \sum_{i=1}^M \sum_{j=1}^N \frac{X(i,j)}{j^2}$	Major Axis Length: It is the length of the major axis in terms of pixels of the ellipse (shape of segmented tumor region)
2	Long Run Emphasis (LRE) measures the distribution of long runs: $LRE = \frac{1}{n_r n_c} \sum_{i=1}^M \sum_{j=1}^N X(i,j) * j^2$	Minor Axis Length: It is the length of the minor axis in terms of pixels of the ellipse (shape of segmented tumor region)
3	Gray-Level Non-uniformity (GLN) measures the similarity of gray level values throughout the image: $GLN = \frac{1}{n_r n_c} \sum_{i=1}^M \sum_{j=1}^N (X(i,j))^2$	Roundness/Circularity(C) $C = \sqrt{\frac{4 * A}{P^2}}$
4	Run Length Non-uniformity (RLN) measures the similarity of the length of runs throughout the image: $RLN = \frac{1}{n_r n_c} \sum_{i=1}^M \sum_{j=1}^N (X(i,j))^2$	Eccentricity (Ecc) $Ecc = \frac{\sqrt{Major\ axis^2 + Minor\ axis^2}}{Major\ axis}$
5	Run Percentage (RP) measures the homogeneity and the distribution of runs in a specific direction: $RP = \frac{n_r n_c}{X(i,j) * j}$	Equiv Diameter (D) $D = \sqrt{\frac{4 * A}{\pi}}$
6	Low Gray-level Run Emphasis (LGRE) measures the distribution of low-gray-level-values: $LGRE = \frac{1}{n_r n_c} \sum_{i=1}^M \sum_{j=1}^N \frac{X(i,j)}{i^2}$	‘A’ represents area of the segmented region. ‘P’ represents perimeter of the segmented region.
7	High Gray-level Run Emphasis (HGRE) measures the distribution of high-gray-level-values: $HGRE = \frac{1}{n_r n_c} \sum_{i=1}^M \sum_{j=1}^N X(i,j) * i^2$	
8	Short Run Low Gray-level Emphasis (SRLGE) measures the joint distribution of short runs and low gray level values: $SRLGE = \frac{1}{n_r n_c} \sum_{i=1}^M \sum_{j=1}^N \frac{X(i,j)}{i^2 * j^2}$	
9	Short Run High Gray-level Emphasis (SRHGE) measures the joint distribution of short runs and high gray level values: $SRHGE = \frac{1}{n_r n_c} \sum_{i=1}^M \sum_{j=1}^N \frac{X(i,j) * j^2}{i^2}$	
10	Long Run Low Gray-level Emphasis (LRLGE) measures the joint distribution of long runs and low gray level values: $LRLGE = \frac{1}{n_r n_c} \sum_{i=1}^M \sum_{j=1}^N \frac{X(i,j) * i^2}{j^2}$	
11	Long Run High Gray-level Emphasis (LRHGE) measures the joint distribution of long runs and high gray level values: $LRHGE = \frac{1}{n_r n_c} \sum_{i=1}^M \sum_{j=1}^N X(i,j) * i^2 * j^2$	

Now the LBP is applied on former three sub band images of the segmented Glioma region.

Local Binary Pattern and GLRL: The LBP concept was proposed by Ojala et al. [32], since then LBP it is widely used and often termed as the most efficient feature descriptor for the purpose of texture classification. In LBP, a mask of size 3x3 is employed against the pixels (neighbourhood) to characterize a particular texture and assess a local binary pattern (LBP). The LBP operator generates information which is undeviating from the local greyscale randomness in the image [33]. So it is generated for every pixel surrounding the central pixel g_c by taking the grey level intensities of a circular neighbourhood. The LBP operators for three DWT sub-bands images are

defined by Eq. (8), Eq. (9) and Eq. (10) given as:

$$LBP_{LL}(P, R) = \sum_{p=0}^{p-1} s(g_p - g_c)2^p \quad (8)$$

$$\text{and } s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$LBP_{HL}(P, R) = \sum_{p=0}^{p-1} s(g_p - g_c)2^p \quad (9)$$

$$\text{and } s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$LBP_{LH}(P, R) = \sum_{p=0}^{p-1} s(g_p - g_c)2^p \quad (10)$$

$$\text{and } s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

where, g_c and g_p represents the grey level intensities of central pixel and neighborhood pixels respectively. 2^p is the binomial weight assigned to each signs $(g_p - g_c)2^p$. This transforms into a unique LBP code with the differences of neighborhood. The joint distribution of center pixel value and their differences physically signifies the overall luminance of an image. Afterwards, LBP code is indexed and GLRL matrix [29] in principle four directions i.e. 0° , 45° , 90° , 135° are formed to count occurrences of runs length for each gray level. For each sub-bands, 44 features are generated. Let $X(i, j)$ denotes total number of occurrences of runs length N at gray level M for the image with n_r rows and n_c columns. Based on $X(i, j)$, eleven features given below are calculated [34]. Here, i dimension corresponds to the gray level and has a length equal to the maximum gray level (M) and j dimension corresponds to the run length and has a length equal to the maximum run length (N). A total of 137 features out of which 132 textural features and five morphological features are extracted from the segmented Glioma region. Table 1 given below simply gives the description of all the eleven texture based GLRL features and five morphological features used in this approach. The entire process of proposed EWBPRM method is illustrated with the help of Figure 2 below: (a) Segmented Glioma tumor image, (b) four DWT sub bands (c) LBP image of LL sub band, (d) LBP image of LH sub band, (e) LBP image of HL sub band, (f) A small region from LBP LL gray image, (g) A small region from LBP LH gray image, (h) A small region from LBP HL gray image, (i) Indexed LBP LL image in 8 bins, (j) Indexed LBP LH image in 8 bins, (k) Indexed LBP HL image in 8 bins.

The algorithm of this feature extraction stage is given below:

Input: Segmented glioma tumor region

Output: Total 137 textural features and morphological or shape based features.

Steps:

1. Initially five shape based features are directly calculated from the segmented glioma tumor region
2. Then apply DWT to the segmented glioma tumor region in order to generate four sub bands i.e. LL, LH, HL and HH.
3. Discard the HH DWT sub band as it mostly contains the noise.
4. Generate LBP image of corresponding LL, LH and HL sub bands.
5. Now LBP images of LL, LH and HL sub bands are indexed in 8 bins.
6. Utilizing this indexed LBP_{LL} image, eleven GLRL textural features are calculated in the principle four directions like 0° , 45° , 90° , 135° and resulted in total 44 features values.
7. Now utilizing this indexed LBP_{LH} image, eleven GLRL textural features are calculated in the principle four directions like 0° , 45° , 90° , 135° and resulted in total 44 features values.
8. Now utilizing this indexed LBP_{HL} image, eleven GLRL textural features are calculated in the principle four directions like 0° , 45° , 90° , 135° and resulted in total 44 features values.

Total 132 textural features and five morphological or shape based features are calculated.

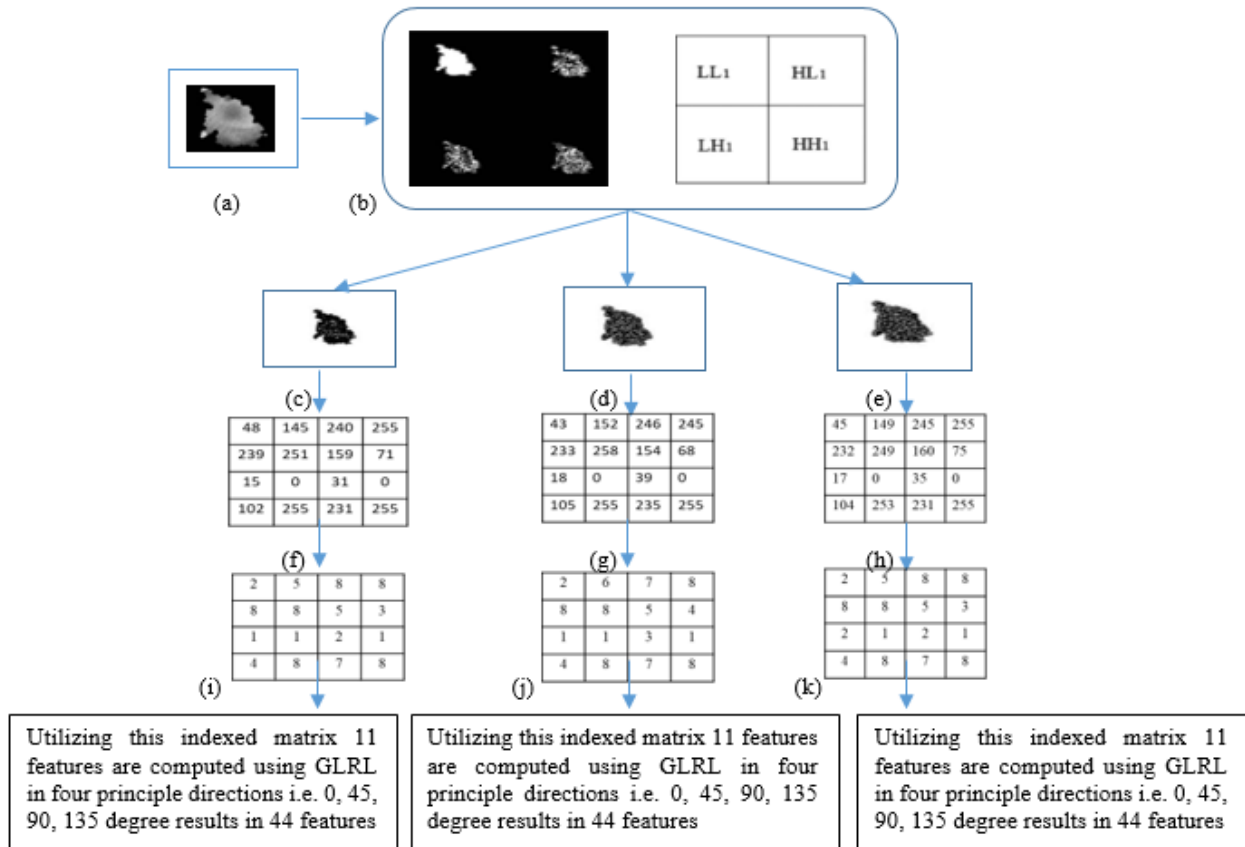


Figure 2. The entire process of texture feature extraction using proposed EWBPRM method

3.3 XGBRF hybrid ensemble learning model

In this paper, XGBoost with Random forest (XGBRF) hybrid ensemble model is used for the classification of Glioma. As XGBoost and Random forest are both popular decision tree based algorithms. XGBoost is an example of boosting algorithm, whereas Random forest is a good example of Begging algorithms. So their combination can be termed as hybrid ensemble learning model. In this model Random forest simply replaces the decision tree as base estimator [35]. The XGBRF is actually an enhanced version of XGBoost classifier. The XGBRF trains random forest instead of gradient boosting decision trees used directly by XGBoost classifier and renders high accuracy on various datasets. The XGBRF simply harness the advantages of both XGBoost and Random forest model in order to offer high accuracy, stability and avoid over fitting problem [36]. Here input to the boosting algorithm is a training dataset D consist of M samples i.e. $D = (x_i, y_i)$ where $i = 1, 2, 3, \dots, M$ and $x_i \in F$ represent feature vector of dataset D and $y_i \in C = \{0, 1\}$ is the corresponding class label. The class label 0 represents low grade Glioma and 1 represents high grade Glioma. As XGBoost is a scalable end-to-end tree boosting system, which include the best tree model into the current classification model, weights play an important role in XGBoost [37]. In it, weights are assigned to all the independent variables, which are then fed into the decision tree and hence ultimately predict the results. Weights of variables predicted wrong by tree is increased and these variables are then fed to the second decision tree. These individual classifiers are the ensemble to give strong and more precise model. Since XGBoost offers regularization which helps in avoid over fitting. XGBoost uses booster = gmtree parameter for classification i.e. tree is grown one after other and attempts to reduce misclassification rate in subsequent iterations. This approach is implemented with the help of Matlab and Python. In this Matlab invokes a python process to use the XGBoost Library provided by Python as XGBoost function is not provided by the Matlab. So this approach is implemented with the help of both Matlab and Python.

Random forest is an enhanced begging algorithm which offers properties like handling of thousands of inputs variables and important feature estimation for classification apart from that it is robust to noise, outliers, simple and parallel in nature. Random forest is a union of classifiers (tree structure) in which each classifier depends on the values of an independently sampled random vector and all the tree in the forest with same distribution [38]. The construction of each tree in the forest from the training dataset requires following points:

1. Trees in the forest should be built from the bootstrap samples. Bootstrap samples are the sample having the same size N as the number of samples in the training set is N.
2. If dataset feature is represented by R then number $r < R$ are defined like at each node r features are opted out of R feature at random.
3. Every tree is grown to the largest extent without pruning.

The algorithm of XGBoost with Random forest (XGBRF) hybrid ensemble model is as follows:

Input: D is a Dataset
 L is the Loss function
 P is the number of iterations
 T is the number of trees
 Θ is a random vector used to construct a tree in the forest
 b(x) is a base learner

Initialize: $f^0(x)$;
 For p = 1 to P do
 For t = 1 to T do
 Generate a vector Θ_t with weight $w_p(i)$
 $S_t \leftarrow$ bootstrap samples(S)
 $C_t \leftarrow$ build tree classifier (S_t, Θ_t)
 Return a hypothesis based on voting
 End for
 Calculate $g_k = \frac{\partial L(y, f)}{\partial f}$;
 Calculate $h_k = \frac{\partial^2 L(y, f)}{\partial f^2}$;
 The structure is determined by selecting splits the maximum gain

$$A = \frac{1}{2} \left[\frac{G_L^2}{H_L} + \frac{G_R^2}{H_R} + \dots \frac{G^2}{H} \right];$$

The leaf weight is determined by $w^* = \frac{G}{H}$;
 Then base learner is determined $b(x) = \sum_{j=1}^T w_j$;
 Add forest $f_k(x) = f_{k-1}(x) + b(x)$;
 End

Result: $f(x) = \sum_{k=0}^M f_k(x)$

During the implementation of XGBRF hybrid ensemble model, the used values of different parameters are illustrated with the help of Table 2 below.

Table 2. Parameters values used in the XGBoost with Random forest (XGBRF) hybrid ensemble model

S.No	Parameters	Values to be taken
1	Number of trees to be used in model	100
2	Learning rate	1
3	Minimum split loss reduction	Logistic
4	Maximum depth of a tree	5
5	Ratio of random samples to be considered for training	9:10
6	Ratio of features to be used to train each tree	1:7

4. RESULTS AND EVALUATION

The Matlab2018a platform is used for the purpose of simulation. The system used for processing is powered by 8th Generation Intel core i5 processor with 8 GB RAM. For the experimentation and effective evaluation of the proposed system, there are basically two types of MRI datasets used in this paper i.e. one local dataset and two global datasets. These datasets are described below:

Local dataset: This dataset comprises of the MRI sequences of low and high grade Glioma tumour gathered from the Sanya Diagnostic centre located at Jabalpur and Bhopal. This dataset consists of T1, T2, T1C and flair MRI images of around 328 patients diagnosed with low grade and high Grade Glioma. All these patients belong to the age group of 20-68 years. For training purpose, 600 high grade Glioma and 600 low grade Glioma MRI images are used from this dataset. Whereas for testing, 200 low grade and 200 high grade MRI images are used. All these MRI images are acquired with the aid of 1.5 GE Health care USA scanners installed in these centres. Ground truth for this local dataset is provided by the chief radiologists of these centres.

BRATS 2013: It is a publically available popular dataset, which is frequently used by the researchers. This dataset contains T1, T2, T1C and flair MRI sequences of around eighty patients. In which, forty five cases of high grade Glioma and thirty five cases of Low grade Glioma in the training section. The testing section consist of leader board and challenge subsets. The leader board subsets consist of 4 low grade Glioma and 21 high grade Glioma cases. Whereas the challenge subset consists of only 10 cases of high grade Glioma, which is used for testing. The ground truth for the BRATS 2013 are available [39]. This dataset is available at smir.ch/BRATS/start2013 for download. This dataset is used in various research studies [19-22, 40].

BRATS 2015: BRATS 2015 is also a publically available large dataset of high grade and low grade Glioma MRI sequences available at smir.ch/BRATS/start2015 for download. This large dataset consists of training and testing subsets. The training subset consist of T1, T2, T1C and flair MRI sequences of total two hundred and seventy four patients. Out of which two hundred twenty cases of high grade Glioma and only fifty four cases of low grade Glioma obtained from the BRATS 2012, BRATS 2013 and the NIH Cancer imaging archive (TCIA). The 70% of this training subset is used for training and rest 30% for testing and validation for our proposed system. The ground truth is provided in the training dataset itself. The testing subset consists of 110 cases of unknown grades, so we haven't used this subset as the grades are not available.

Initially the effective segmentation assessment of this system is carried out using the statistical parameters like Jaccard Index (JI), Dice Similarity Coefficient (DSC) and Hausdroff Distance (HD). The best segmentation performance is achieved using the two MRI fusion combinations i.e. T1C+T2+Flair and T1+T1C+T2+Flair as their JI, DSC and

HD values are better. The Table 3 below depicts the segmentation performance of single MRI sequence and their fusion combinations. As the segmentation accuracy of MRI fusion combinations i.e. T1C+T2+Flair and T1+T1C+T2+Flair are high as compare to the single MRI imaging sequence and other MRI fusion combinations, so these two MRI fusion combinations are taken for the Glioma classification experimentation. The resulted confusion matrices a, b and c of the proposed system, which takes both the MRI fusion combinations i.e. T1C+T2+Flair and T1+T1C+T2+Flair on the local dataset, BRATS 2013 dataset and BRATS 2015 dataset respectively are illustrated with the help of the Table 4 and 5. The used classification rates are listed and defined below:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}, \text{Precision} = \frac{TP}{TP+FP},$$

$$\text{Sensitivity} = \frac{TP}{TP+FN}, \text{Specificity} = \frac{TN}{TN+FP},$$

$$\text{F1 score} = \frac{2TP}{2TP + FP + FN}$$

where, TP = True positive means low grade Glioma was present in the test MRI sequences and correctly classified as low grade Glioma.

TN = True Negative means high grade Glioma was present in the test MRI sequences and also correctly classified as high grade Glioma.

FP = False positive means low grade Glioma was present in the test MRI sequences and wrongly classified as high grade Glioma.

FN = False negative means high grade Glioma was present in the test MRI sequences and wrongly classified as low grade Glioma.

Table 3. Segmentation performance assessment on the local dataset with various MRI imaging sequence fusion combinations.

S. No	MRI Modality and their Fusion combinations	Jaccard Index	Dice Similarity Coefficient	Hausdroff Distance
1	T1	0.85	0.87	16.05
2	T1C	0.89	0.91	17.25
3	T2	0.79	0.89	19.56
4	Flair	0.88	0.90	12.23
5	T1+T1C+T2+FLAIR	0.96	0.96	9.85
6	T1C+T2+FLAIR	0.97	0.98	9.10
7	T1+T2+FLAIR	0.91	0.90	12.85
8	T1C+FLAIR	0.84	0.89	14.21
9	T1+T2	0.65	0.68	19.25
10	T1+T1C	0.56	0.58	21.29

Table 4. Confusion matrices of the proposed approach based on the fusion of three T1C+T2+Flair MRI modalities

LGG=200 HGG=200	Low grade Glioma	High grade Glioma	LGG=04 HGG=31	Low grade Glioma	High grade Glioma	LGG=18 HGG=66	Low grade Glioma	High grade Glioma
Low grade Glioma	198	02	Low grade Glioma	04	00	Low grade Glioma	15	03
High grade Glioma	01	199	High grade Glioma	01	30	High grade Glioma	01	65

(a)

(b)

(c)

Table 5. Confusion matrices of the proposed approach based on the fusion of all four T1+T1C+T2+Flair MRI modalities

LGG=200 HGG=200	Low grade Glioma	High grade Glioma	LGG=04 HGG=31	Low grade Glioma	High grade Glioma	LGG=18 HGG=66	Low grade Glioma	High grade Glioma
Low grade Glioma	189	11	Low grade Glioma	03	01	Low grade Glioma	14	04
High grade Glioma	10	190	High grade Glioma	04	27	High grade Glioma	04	62

(a)

(b)

(c)

Table 6. Proposed approach classification rates on the local dataset

MRI fusion combination	Accuracy	Precision	Sensitivity	Specificity	F1 score
T1+T1C+T2+Flair	96.75	96	97.46	96.06	96.73
T1C+T2+Flair	99.25	99	99.50	99	99.25

Table 7. Proposed approach classification rates on global Dataset (BRATS 2013)

MRI fusion combination	Accuracy	Precision	Sensitivity	Specificity	F1 score
T1+T1C+T2+Flair	91.43	100	60	100	72.73
T1C+T2+Flair	97.22	100	80	100	88.89

Table 8. Proposed approach classification rates on global Dataset (BRATS 2015)

MRI fusion combination	Accuracy	Precision	Sensitivity	Specificity	F1 score
T1+T1C+T2+Flair	90.48	77.78	77.78	93.94	77.78
T1C+T2+Flair	95.24	83.3	93.75	95.59	88.27

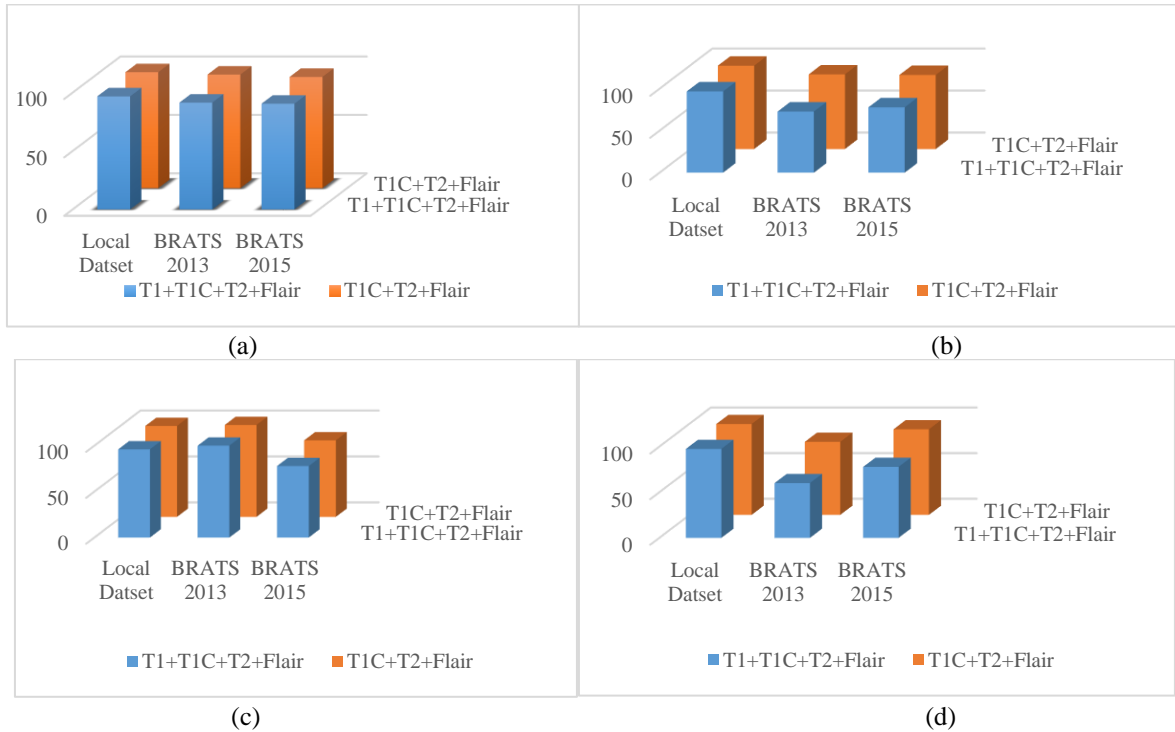


Figure 3. Comparison graphs of the proposed approach with both the MRI fusion combinations on all the three datasets

Table 9. Classification rates of proposed EWPRL method with different popular classifiers using TIC+T2+Flair

Combinations on Local Dataset	Accuracy	Precision	Sensitivity	Specificity	F1 score
(DWT+LBP+GLRL based texture and shape based features)+SVM	92.5	91	93.81	91.26	92.39
(DWT+LBP+GLRL based texture and shape based features)+NB	95.25	95	95.4	95.2	95.24
(DWT+LBP+GLRL based texture and shape based features)+RF	96	95.7	96	95.5	95.7
(DWT+LBP+GLRL based texture and shape based features)+KNN	95	94.5	95.45	94.55	94.97
(DWT+LBP+GLRL based texture and shape based features)+ADBRF	97	96.5	97.7	96.55	96.5
(DWT+LBP+GLRL based texture and shape based features)+XGBRF	99.25	99	99.50	99	99.25

Table 10. Classification rates of proposed EWPRL method along with different popular classifiers using T1+T1C+T2+Flair

Combinations on Local Dataset	Accuracy	Precision	Sensitivity	Specificity	F1 score
(DWT+LBP+GLRL based texture and shape based features)+SVM	89.5	89	89.9	89	89.45
(DWT+LBP+GLRL based texture and shape based features)+NB	93	95	91.35	94.79	93.14
(DWT+LBP+GLRL based texture and shape based features)+RF	94.25	95.5	93.17	95.3	94.32
(DWT+LBP+GLRL based texture and shape based features)+KNN	91	94	88.68	93.6	91.26
(DWT+LBP+GLRL based texture and shape based features)+ADBRF	95.25	96	94.58	95.94	95.29
(DWT+LBP+GLRL based texture and shape based features)+XGBRF	96.75	96	97.46	96.06	96.73

Table 11. Performance comparison of the proposed system with the existing approaches over the local dataset

Author and year	Subashini et al. 2015	Zhan et al. 2017	Gupta et al. 2018	Gupta et al. 2019	Proposed approach
Feature extraction method and Classifier used	GLCM+SVM	LBP+KNN	GLRL+NB	LBP+GLRL, GLCM + Ensemble (KNN+NB+SVM)	EWBPRL + Hybrid ensemble learning classifier (XGBRF)
Accuracy	93	87.75	95	97.5	99.25
Precision	92	89.5	95.5	98	99
Sensitivity	93.88	86.47	94.55	97.03	99.50
Specificity	92.16	89.12	95.45	97.98	99
F1 Score	92.93	87.96	95.5	97.51	99.25
Processing or training time (seconds)	86	98	106	198	82

Table 12. Performance comparison of the proposed system with the existing approaches over the BRATS 2013 and 2015 datasets

Author and year	Subashini et al. 2015		Zhan et al. 2017		Gupta et al. 2018		Gupta et al. 2019		Proposed approach	
Feature extraction method and Classifier used	GLCM+SVM		LBP+KNN		GLRL+NB		LBP+GLRL, GLCM + Ensemble (KNN+NB+SVM)		EWBPRL + Hybrid ensemble learning classifier (XGBRF)	
BRATS Glioma Datasets	2013	2015	2013	2015	2013	2015	2013	2015	2013	2015
Accuracy	91.25	85.26	85.25	90.1	92.98	90.5	94.25	92.98	97.22	95.24
Precision	90.12	75.23	86	76.2	93.89	76.26	94.98	78.26	100	83.3
Sensitivity	74.75	84.25	76.25	86.4	78.99	87.12	78.85	89.25	80	93.75
Specificity	90.24	89.25	87.99	88.98	93.25	90.2	94.98	90.12	100	95.59
F1 Score	80.12	78.25	85.45	80.45	86.98	81.45	86.21	82.2	90	88.27
Processing or training time (seconds)	74	80	78	85	84	90	156	165	75	78

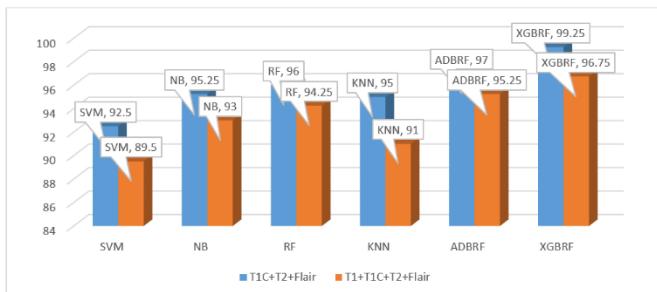


Figure 4. Accuracy comparison graph of the proposed EWBPRL method along with different popular classifiers

The above confusion matrixes simply illustrate the performance of the proposed system for two fusion combinations of MRI modalities i.e. T1+T1C+T2+Flair and T1C+T2+Flair on the local, BRATS 2013 and BRATS 2015 datasets. Based on these, above mention classification rates are measured and summarized with the help of Table 6, 7 and 8. The Table 6 simply presents the different classification rates on the local dataset. Whereas Table 7 and 8 presents classification rates on the BRATS 2013 and 2015 datasets. On the local dataset, the observed accuracy is 99.25%. The Figure 3 above also presents the comparison graph of the proposed system with both the MRI fusion combinations on all the three datasets in terms of Accuracy (a), F1 score (b), Precision (c) and Sensitivity (d). The segmentation and proposed feature extraction approach i.e. EWBPRL method is also used along with various popular classifiers on the likes of SVM (Support

vector machine), Random Forest (RF), Naïve Bayes (NB), K-Nearest Neighbor (KNN) and ADaBoost with Random Forest (ADBRF) classifier [44]. Their performance in terms of classification rates on the local dataset are summarized with the help of Table 9 for fused MRI sequence of TIC+T2+Flair and Table 10 for fused MRI sequence of T1+TIC+T2+Flair above. Both the MRI fusion combinations are used along with these classifiers. Figure 4 simply illustrates an accuracy comparison graph among the above mention classifiers using both fused MRI sequences combinations like TIC+T2+Flair and T1+TIC+T2+Flair on the local Dataset.

Finally the proposed approach is compared with the already existing state of the art approaches of Glioma classification based on machine learning and ensemble learning. All these approaches are implemented and validated in the same computing environment. The Table 11 and 12 simply depicts the performance comparison of the proposed system with the existing or already proposed approaches on the local and global datasets i.e. BRATS 2013 & 2015. This comparison is done with the help various classification rates which are defined above. It is evident from the Table 11 and 12 that the proposed approach outperforms the already existing system or approaches in terms of accuracy, Precision, Sensitivity, Specificity, F1 score and computational time. The proposed system achieves an accuracy of 99.25%, which is clearly higher as compared to the already existing approaches. The screenshot of the proposed working automated system for the Glioma segmentation and classification with result is depicted with the help of Figure 5.

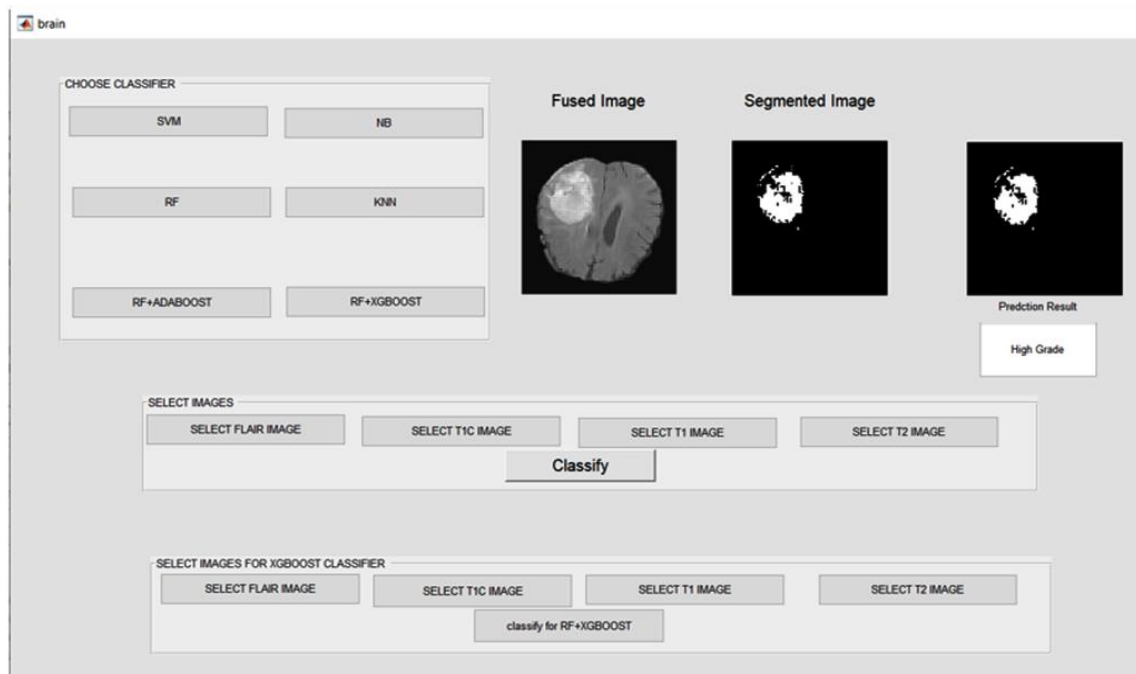


Figure 5. Screenshot of the working automated system for the Glioma segmentation and classification with result

5. DISCUSSION

The proposed system for Glioma classification offers best results while using the T1C+T2+Flair MRI sequence combination. The segmentation performance illustrated with the help of Table 3 clearly showcased that the two MRI sequence combinations (T1C+T2+Flair and T1+T1C+T2+Flair) offers best results. So these two MRI sequence combinations are used in this paper for the sake of experimentation and evaluation. Whereas most of the state of the art approaches for Glioma classification has only either used single MRI modality like Flair, T2 or used T1+T1C+T2+Flair MRI sequence combination. This paper has explored and established the fact that T1C+T2+Flair MRI sequence combination offers highest accuracy and will deliver best results in real time. This fact is well proved with the help of the confusion matrices and classification rates values represented by Table 4, 5, 6, 7, and 8. These classification rates are obtained by properly tuning the XGBRF ensemble classifier. This classifier is used for the first time for any brain tumor or cancer type classification. The XGBRF parameter value especially number of trees to be used has great effect over the performance of this Glioma classification system and tends to provide optimum accuracy at 100 value. The performance of the system especially validation accuracy tends to decrease while exceeding this 100 value. These MRI fusion combinations along with the segmentation and proposed feature extraction method i.e. EWPRL is used along with other popular machine learning classifiers like SVM, NB, KNN, RF and ADBRF. This experimentation is done in order to check effectiveness, adaptability and robustness of the two MRI fusion combinations segmentation and proposed texture feature extraction method. The performance of all these classifiers is very encouraging and represented by the Table 9, 10. The ADBRF classifier along with MRI fusion combinations segmentation in this paper (T1C+T2+Flair and T1+T1C+T2+Flair) and proposed texture feature extraction method was very good and close to the

performance of XGBRF classifier. The existing state of the art approaches for Glioma classification was compared with the proposed system on the highly balance local dataset as well as on the unbalance global datasets like BRATS 2013 and 2015 using various classification parameters. This comparison is done in order to prove the superiority and effectiveness of the proposed system based on proposed EWPRL method and XGBRF hybrid ensemble classifier model for Glioma classification. The Table 11 and 12 simply proves that the proposed Glioma classification system is better in terms of delivering accurate results as well as efficient in terms of computational or training time as compare to the existing state of the art approaches.

6. CONCLUSION

The proposed system for Glioma classification is highly accurate as per the classification rates observed on all the three datasets i.e. local dataset, BRATS 2013 and BRATS 2015 dataset. This system offers an accuracy of 99.25% on the balance local dataset with T1C+T2+Flair MRI imaging fusion combination. Whereas this system also delivers good performance on the unbalance global datasets with an accuracy of 97.22 and 95.24 on the BRATS 2013, 2015 datasets. The proposed system also takes less training time as compare to other state of the art approaches based on SVM, KNN, NB and ensemble classifiers. The proposed texture feature extraction method and use of XGBRF ensemble classifier has greatly enhanced the performance of this system. The proposed feature extraction method allows texture features to be extracted at multi-resolutions at both frequency and time domain. A total of 132 texture features and 5 morphological features are extracted, which proved to be very effective in terms of accurate Glioma classification. The XGBRF ensemble classifier based on XGBoost and Random forest is used for the first time for any brain tumor or any cancer type classification delivers very good results. These

two later stages are mainly responsible for the good performance of this proposed Glioma classification system. In future, this work can be extended for the multiclass classification of other types of brain tumors like Central nervous system, secondary and primary brain tumors.

ACKNOWLEDGMENTS

We like to thank Dr. Pushpraj Bhatele, Chief radiologist MP MRI and CT scan centre at Netaji Subhash Chandra Bose Medical College, Jabalpur, Sanya Diagnostic MRI center, Bhopal and Cancer Hospital Gwalior for providing the clinical MRI dataset of brain MRI images. I am thankful to all the radiologists for providing their valuable support in terms of providing data, knowledge and validate this proposed work.

REFERENCES

- [1] American Cancer Society, Cancer Facts & Figures 2015. Technical report, American Cancer Society, Georgia, 2015.
- [2] Over 2,500 Indian kids suffer from brain tumor every year. The Hindu. [https://www.thehindu.com/scitech/health/Over-2500-Indian-kids-suffer-from-brain-tumour every year/article14418512.ece/2016](https://www.thehindu.com/scitech/health/Over-2500-Indian-kids-suffer-from-brain-tumour%20every%20year/article14418512.ece/2016), accessed on 26 August 2019.
- [3] What is a Brain tumor? The brain tumor charity blog. <https://www.thebraintumourcharity.org/understanding-brain-tumours/symptoms-andinformation/what-is-a-brain-tumor/>, accessed on 25 August 2019.
- [4] OncoLink Team: All about Pediatric Gliomas (Low and High Grade). <https://www.oncolink.org/cancers/braintumors/all-about-pediatric-Gliomas-low-and-high-grade> (2018). Accessed on 26 August 2019.
- [5] Ain, Q., Jaffar, M.A., Choi, T.S. (2014). Fuzzy anisotropic diffusion based segmentation and texture based ensemble classification of brain tumor. *Applied Soft Computing*, 21: 330-340. <https://doi.org/10.1016/j.asoc.2014.03.019>
- [6] Boudraa, A.O., Dehak, S.M.R., Zhu, Y.M., Pachai, C., Bao, Y.G., Grimaud, J. (2000). Automated segmentation of multiple sclerosis lesions in multispectral MR imaging using fuzzy clustering. *Computers in Biology and Medicine*, 30(1): 23-40. [https://doi.org/10.1016/S0010-4825\(99\)00019-0](https://doi.org/10.1016/S0010-4825(99)00019-0)
- [7] Van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A., Suetens, P. (2001). Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE Transactions on Medical Imaging*, 20(8): 677-688. <https://doi.org/10.1109/42.938237>
- [8] Wang, Z., Hu, Q., Loe, K., Aziz, A., Nowinski, W.L. (2004). Rapid and automatic detection of brain tumors in MR images. *Proceedings of SPIE Medical Imaging*, San Diego. <https://doi.org/10.1117/12.538035>
- [9] Ray, N., Saha, B.N., Brown, M. (2007). Locating brain tumors from MR imagery using symmetry. 2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, pp. 224-228. <https://doi.org/10.1109/ACSSC.2007.4487200>
- [10] Khayati, R., Vafadust, M., Towhidkhal, F., Nabavi, S.M. (2008). A novel method for automatic determination of different stages of multiple sclerosis lesions in brain MR FLAIR images. *Computerized Medical Imaging Graphics*, 32(2): 124-133. <https://doi.org/10.1016/j.compmedimag.2007.10.003>
- [11] Wang, H., Fei, B. (2009). A modified fuzzy C-means classification method using a multiscale diffusion filtering scheme. *Medical Image Analysis*, 13(2): 193-202. <https://doi.org/10.1016/j.media.2008.06.014>
- [12] Yamamoto, D., Arimura, H., Kakeda, S., Magome, T., Yamashita, Y., Toyofuku, F., Ohki, M., Higashida, Y., Korogi, Y. (2010). Computer-aided detection of multiple sclerosis lesions in brain magnetic resonance images: False positive reduction scheme consisted of rule-based level set method, and support vector machine. *Computerized Medical Imaging Graphics*, 34(5): 404-413. <https://doi.org/10.1016/j.compmedimag.2010.02.001>
- [13] Sharma, M., Mukherjee, S. (2013). Fuzzy C-means and snake model for segmenting astrocytoma - a type of brain tumor. *Int. J. Adv. Eng. Sci.*, 3(3): 30-35.
- [14] El-Dahshan, E.A., Hosny, T., Salem, A.M. (2010). Hybrid intelligent techniques for MRI brain images classification. *Digital Signal Processing*, 20(2): 433-441. <https://doi.org/10.1016/j.dsp.2009.07.002>
- [15] El-Dahshan, E.A., Mohsen, H.M., Revett, K., Salem, A.M. (2014). Computer-aided diagnosis of human brain tumor through MRI: A survey and anew algorithm. *Expert Systems with Applications*, 41(11): 5526-5545. <https://doi.org/10.1016/j.eswa.2014.01.021>
- [16] Arakeri, M.P., Reddy, G.R.M. (2015). Computer-aided diagnosis system for tissue characterization of brain tumor on magnetic resonance images. *Signal Image Video Processing*, 9: 409-425. <https://doi.org/10.1007/s11760-013-0456-z>
- [17] Gupta, N., Khanna, P. (2013). A fast and efficient computer aided diagnostic system to detect tumor from brain magnetic resonance imaging. *Int. J. Imaging Syst. Tech.*, 25(2): 123-130. <https://doi.org/10.1002/ima.22128>
- [18] Subashini, M.M., Sahoo, S.K., Sunil, V., Easwaran, S. (2016). A non-invasive methodology for the grade identification of astrocytoma using image processing and artificial intelligence techniques. *Expert System with Applications*, 43: 186-196. <https://doi.org/10.1016/j.eswa.2015.08.036>
- [19] Vishnuvarthanan, G., Rajasekaran, M.P., Subbaraj, P., Vishnuvarthanan, A. (2016). An unsupervised learning method with a clustering approach for tumor identification and tissue segmentation in magnetic resonance brain images. *Applied Soft Computing*, 38: 190-212. <https://doi.org/10.1016/j.asoc.2015.09.016>
- [20] Zhan, T., Feng, P., Hong, X., Lu, Z., Xiao, L., Zhang, Y. (2017). An automatic Glioma grading method based on multi-feature extraction and fusion. *Technology and Health Care*, 25(S1): 377-385. <https://doi.org/10.3233/THC-171341>
- [21] Gupta, N., Bhatele, P., Khanna, P. (2018). Identification of Gliomas from brain MRI through adaptive segmentation and run length of centralized patterns. *Journal of Computational Science*, 25: 213-220. <https://doi.org/10.1016/j.jocs.2017.02.009>
- [22] Gupta, N., Bhatele, P., Khanna, P. (2019). Glioma detection on brain MRIs using texture and morphological

- features with ensemble learning. *Journal of Biomedical Signal Processing and Control*, 47: 115-125. <https://doi.org/10.1016/j.bspc.2018.06.003>
- [23] Golshan, H.M., Hasanzedeh, R.P.R., Yousefzadeh, S.C. (2013). An MRI de-noising method using data redundancy and local SNR estimation. *Magnetic Resonance Imaging*, 31(7): 1206-1217. <https://doi.org/10.1016/j.mri.2013.04.004>
- [24] Davies, E.R. (2018). *The Role of Thresholding. Computer Vision (Fifth Edition)*. Academic Press, Science Direct, pp. 93-118.
- [25] Upadhyay, N., Waldman, A.D. (2011). Conventional MRI evaluation of gliomas. *The British Journal of Radiology*, 84(2): 107-S111. <https://doi.org/10.1259/bjr/65711810>
- [26] Chaddad, A., Desrosiers, C., Hassan, L., Tanougast, C. (2016). A quantitative study of shape descriptors from glioblastoma multiforme phenotypes for predicting survival outcome. *The British Journal of Radiology*, 89(1068). <https://doi.org/10.1259/bjr.20160575>
- [27] Hu, L.S., Ning, S., Eschbacher, J.M., Gaw, N., Dueck, A.C., Smith, K.A., Nakaji, P., Plasencia, J., Plasencia, S., Price, S.J., Tran, N., Loftus, J., Jenkins, R., O'Neill, B.P., Elmquist, W., Baxter, L.C., Gao, F., Frakes, D., Karis, J.P., Zwart, C., Swanson, K.R., Sarkaria, J., Wu, T., Mitchell, J.R., Li, J. (2015). Multi-parametric MRI and texture analysis to visualize spatial histologic heterogeneity and tumor extent in glioblastoma. *PLOS ONE*, 10(11): e0141506. <https://doi.org/10.1371/journal.pone.0141506>
- [28] Liu, L., Fieguth, P., Kuang, G. (2011). Generalized local binary patterns for texture classification. *Proceedings of Br. Mach. Vis. Conference (BMVC)*, pp. 1-11.
- [29] Loh, H.H., Leu, J.G., Luo, R.C. (1998). The analysis of natural textures using run length features. *IEEE Transactions on Industrial Electronics*, 35(2): 323-328. <https://doi.org/10.1109/41.192665>
- [30] Gonzalez, R.C., Eddins, S.L., Woods, R.E. (2004). *Digital Image Publishing Using MATLAB*. Englewood Cliffs, Prentice Hall.
- [31] Yu, M., Deng, K., Yang, H., Qin, C. (2018). Improved WasH feature matching based on 2D-DWT for stereo remote sensing images. *Sensors (Basel) MDPI*, 18(10): 3494. <https://doi.org/10.3390/s18103494>
- [32] Ojala, T., Pietikainen, M., Harwood, D. (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions in Pattern Recognition. *Proceedings of 12th International Conference on Pattern Recognition*, Jerusalem, Israel, pp. 582-585. <https://doi.org/10.1109/ICPR.1994.576366>
- [33] Ojala, T., Pietikainen, M., Maenpaa, T. (2002). Multi-resolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7): 971-987. <https://doi.org/10.1109/TPAMI.2002.1017623>
- [34] Tang, X. (1998). Texture information in run-length matrices. *IEEE Transactions on Image Processing*, 7(11): 1602-1609. <https://doi.org/10.1109/83.725367>
- [35] Brandl, G. (2019). Get Started with XGBoost. <https://xgboost.readthedocs.io/en/latest/tutorials/rf.html>, accessed on 8 October 2019.
- [36] Dong, X., Yu, Z., Cao, W., Shi, Y., Ma, Q. (2020). A survey on ensemble learning. *Frontiers of Computer Science*, 14: 241-258. <https://doi.org/10.1007/s11704-019-8208-z>
- [37] Chen, T., Guestrin, C. (2016). XGBoost a scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*, New York, pp. 785-794.
- [38] Breiman, L. (2001). Random forest. *Machine Learning*, 45: 5-32. <https://doi.org/10.1023/A:1010933404324>
- [39] Kistler, M., Bonaretti, S., Pfahrer, M., Niklaus, R., Büchler, P. (2013). The virtual skeleton database: an open access repository for biomedical research and collaboration. *Journal of Medical Internet Research*, 15(11): e245. <https://doi.org/10.2196/jmir.2930>
- [40] Hambali, M., Saheed, Y., Oladele, T., Gbolagade, M. (2019). ADaBoost ensemble algorithms for breast cancer classification. *Journal of Advances in Computer Research*, 10(2): 31-52.